in a manner that provides enhanced performance over conventional storage system designs not possessing these capabilities.

In one respect then, disclosed herein is a resource model that takes into account I/O resources such as disk drive capacity and/or memory availability. The resource model may be capable of estimating information management system I/O resource utilization. The resource model may also be used, for example, by a resource manager to make decisions on whether or not a system is capable of supporting additional clients or viewers, and/or to adaptively change read-ahead strategy so that system resource utilization may be balanced and/or optimized. The resource model may be further capable of discovering a limitation on read-ahead buffer size under exceptional conditions, *e.g.,* when client access pattern is highly skewed. A limit or cap on read-ahead buffer size may be further incorporated so that buffer memory resource may be better utilized. In one embodiment, the resource model may incorporate an algorithm that considers system design and implementation factors in a manner so that the algorithm is capable of yielding results that reflect actual system dynamics.

In another respect, disclosed herein is a resource model that may be used, for example, by a resource manager, to modify the cycle time of one or more storage devices of an unequally-loaded multiple storage device system in order to better allocate buffer space among unequally-loaded storage devices (*e.g.,* multiple storage devices containing content of different popularity levels). In this regard, read-ahead may become unequal when disk access pattern (*i.e.,* workload) is unequal. This capability may be implemented, for example, to lower the cycle time of a lightly-loaded disk drive (*e.g.,* containing relatively less popular content) so that the lightly-loaded disk drive uses more input output operations per second ("IOPS"), and consumes less buffer space, thus freeing up more buffer space for use by cache memory and/or for access use by a more heavily-loaded disk drive (*e.g.,* containing relatively more popular content) of the same system. The resource model may also be capable of adjusting cache memory size to optimize system performance, *e.g.,* by increasing cache memory size rather than buffer memory size in those cases where increasing buffer size will result in no throughput improvement.

In yet another respect, disclosed herein is a disk workload monitor that is capable of dynamically monitoring or tracking disk access load at the logical volume level. The disk workload monitor may be capable of feeding or otherwise communicating the monitored disk access load information back to a resource model.

In yet another respect, disclosed is a disk capacity monitor that is capable of dynamically monitoring or measuring disk drive capacity. The disk capacity monitor may be capable of feeding or otherwise communicating the monitored disk drive capacity information back to a resource model.

In yet another respect, disclosed herein is a storage management processing engine that is capable of monitoring system I/O resource workload distribution and/or of detecting workload skew. The monitored workload distribution and/or workload skew information may be fed back or otherwise communicated to an I/O manager or I/O admission controller (*e.g.*, I/O admission control algorithm running in a resource manager of the storage processing engine), and may be taken into account or otherwise considered when making decisions regarding the admission of new requests for information (*e.g.*, requests for streaming content). The disclosed methods and systems may include storage management processing engine software designed for implementation in a resource manager and/or logical volume manager of the storage management processing engine.

In yet another respect, disclosed herein is a resource management architecture that may include a resource manager, a resource model, a disk workload monitor and/or a disk capacity monitor. The resource manager and/or resource model may be in communication with at least one of the resource manager, resource model, or a combination thereof. In operation, monitored workload and/or disk capacity information may be dynamically fed back directly to the resource model or indirectly to the resource model through the resource manager. The resource model may be used by the resource manager to generate system performance information, such as system utilization information, based on the monitored workload and/or storage device capacity information. The resource manager may be configured to use the generated system performance information to perform admission control (*e.g.*, so that the resource manager effectively monitors

workload distribution among all storage devices under its control and uses this information for I/O admission control for the information management system) and/or to advise or instruct the information management system regarding read-ahead strategy.

In one embodiment, the disclosed resource management architecture may be employed to manage delivery of information from storage devices that are capable of performing resource management and I/O demand scheduling at the logical volume level. In this embodiment, read-ahead size or length may be estimated based on designated I/O capacity and buffer memory size, and the method may be used as an admission control policy when accepting new I/O demands at the logical volume level. In one implementation, this embodiment of the disclosed method may be employed to decision whether or not there is enough I/O capacity and buffer memory to support a new viewer's demand for a video object (*e.g.* a movie), and if so, what is the optimal read-ahead size for each viewer that is served by I/O operations based on the available I/O capacity, the buffer memory size, and information related to characteristics of the existing viewers.

In yet another respect, disclosed herein are substantially lightweight or low-processing-overhead methods and systems that may be implemented to support Internet streaming (*e.g.,* including video-on-demand ("VOD") applications). These disclosed methods and systems may utilize workload monitoring algorithms implemented in the storage processor, may further include and consider workload distribution information in I/O admission control calculations/decisions, and/or may further include a lightweight IOPS validation algorithm that may be used to verify system I/O performance characteristics such as "average access time" and "transfer rate" when a system is turned on or rebooted.

In yet another respect, disclosed herein is a network processing system operable to process information communicated via a network environment. The system may include a network processor operable to process network-communicated information and a storage management processing engine operable to perform the I/O resource management features described herein.